# An effective MLP model for detecting malicious nodes in PoS permissionless blockchains

*Njoku ThankGod* Anthony[1*], *Mahmoud* Shafik[1], and *Hany F.* Atlam[2]

[1]College of Science and Engineering, University of Derby, Derby, DE22 3AW, UK
[2]Cyber Security Centre, WMG, University of Warwick, Coventry, CV4 7AL, UK

**Abstract.** With the proliferation of blockchain technology, ensuring the security and integrity of permissionless Proof-of-Stake (PoS) blockchain networks has become imperative. This paper addresses the persistent need for an effective system to detect and mitigate malicious nodes in such environments. Leveraging Deep Learning (DL) techniques, specifically Multi-Layer Perceptron (MLP), a novel model is proposed for real-time identification and detection of malicious nodes in PoS blockchain networks. The model integrates components for data collection, feature extraction, and model training using MLP. The proposed model is trained on labelled data representing both benign and malicious node activities, utilising transaction volumes, frequencies, timestamps, and node reputation scores to identify anomalous behaviour indicative of malicious activity. The experimental results validate the efficacy of the proposed model in distinguishing between normal and malicious nodes within blockchain networks. The model demonstrates exceptional performance in classification tasks with an accuracy of 99%, precision, recall, and F1-score values hovering around 0.99 for both classes. The experimental results verify the proposed model as a dependable tool for enhancing the security and integrity of PoS blockchain networks, offering superior performance in real-time detection and mitigation of malicious activities.

## 1 Introduction

The utilisation of blockchain technology's inherent properties is crucial in identifying rogue nodes within Proof of Stake (PoS) permissionless blockchains. When coupled with resilient consensus algorithms, the decentralised and peer-to-peer framework inherent in blockchain technology holds the promise of fortifying cyber defence capabilities by efficiently eliminating harmful operations [1-2]. In addition, using authentication-based methods using cryptographic certificates in blockchain can be advantageous in identifying malicious and malfunctioning nodes [3]. Various researchers have suggested several methodologies for identifying malicious nodes in the blockchain network. For instance, a data-mining strategy that integrates local node and neighbour node data has been employed to detect intrusions and highlight nodes with malicious intent [4]. Furthermore, a proposed architecture for dynamic trust has been put up to identify self-interested and malicious nodes, hence

* Corresponding author: t.njoku@derby.ac.uk

improving the usefulness of the network [5]. Research in wireless sensor networks has shown that trust management can be strengthened by quickly detecting rogue nodes in large network traffic situations [6].

Furthermore, the incorporation of blockchain technology can enhance the protection and confidentiality of electronic health record systems by utilising characteristics such as selective data sharing, traceability of malicious activities through unchangeable records, and the use of unique public keys for transactions to ensure personal identity [7]. Moreover, the core principle of Proof of Stake (PoS) centres on nodes exhibiting possession of a predetermined quantity of coins, thereby establishing their legitimacy and expertise within the blockchain framework [8]. There have been proposals for trust models that utilise blockchain technology to identify rogue nodes in wireless sensor networks. These models highlight the significant importance of trust in safeguarding network security [9]. Blockchain-based collaborative intrusion detection systems have demonstrated potential in quickly detecting rogue nodes, especially in situations involving pollution attacks [10].

The identification of malicious nodes in PoS permissionless blockchains requires a comprehensive strategy that combines the advantages of blockchain technology, consensus algorithms, authentication-based methods, data mining models, and dynamic trust frameworks. It is conceivable to enhance the security and integrity of blockchain networks against hostile actors by utilising these techniques. This paper proposes a Multi-Layer Perceptron (MLP) model to improve the security and integrity of PoS blockchain networks by focusing on real-time detection and decision-making, scalability considerations, and security measures, which can in practice help protect against malicious actors. The rest of this paper is organised as follows: Section 2 presents the Investigation into the current State of the Art of related work, Section 3 introduces the proposed MLP model, Section 4 discusses the testing, validation, and results, and Section 5 is the conclusion.

## 2 Investigation into the current state of the art of related work

Several researchers have investigated intrusion detection in blockchain platforms. Mansour [11] introduced a new method called PRO-DLBIDCPS for detecting intrusions in Cyber-Physical Systems (CPS) using blockchain technology. The proposed technique incorporates an Adaptive Harmony Search Algorithm (AHSA) to enhance the feature selection, hence enhancing the identification of pertinent feature subsets. The proposed technique also utilised an Attention-based Bi-directional Gated Recurrent Neural Network (ABi-GRNN) model for intrusion detection and classification. The utilisation of the Poor and Rich Optimization (PRO) algorithm-based hyperparameter optimiser significantly enhances the efficiency of ABi-GRNN, leading to improved outcomes in intrusion detection. Furthermore, the utilisation of blockchain technology is employed to augment security within CPS environment. Hisham et al. [12] provided a comprehensive review of the incorporation of anomaly detection models into blockchain technology and evaluated the effectiveness of supervised and unsupervised Machine Learning (ML) techniques to identify both fraudulent and valid transactions. Based on this review, it can be concluded that supervised learning is the predominant approach employed in the investigation of anomalous detection inside blockchain networks.

Li et al. [13] introduce DLBC, a novel approach harnessing the computational power of miners for deep learning training, thereby redefining proof of useful work beyond conventional hash value calculations. The proposed approach overcomes inherent limitations of existing proof of useful work mechanisms by accommodating multiple tasks, larger models, and training datasets and introducing a comprehensive ranking mechanism based on task difficulty factors. The authors enhance the robustness of DLBC using DNN-watermarking techniques. The results suggest that DLBC holds promise for enhancing the security and efficiency of deep learning training on permissionless blockchains.

Baig et al. [14] proposed a Blockchain Ensemble stacked Machine Learning (BEML) approach to enhance the security of IoT networks. Traditional security mechanisms for IoT networks are often vulnerable due to centralisation or reliance on third parties. Moreover, ML-based approaches for attack detection have limitations in accuracy and performance. The BEML approach integrates blockchain, InterPlanetary File System (IPFS), and a stacked ML model for attack detection. In the BEML approach, the blockchain module registers and authenticates network nodes securely, while IPFS stores data with unique hashes for access. The ML model, comprising multiple single learner algorithms, is combined to compensate for individual weaknesses, resulting in improved detection accuracy. This model is used to detect Denial of Service (DoS) attacks and identify malicious nodes for removal from the network. Simulation results demonstrate the efficiency and effectiveness of the proposed approach in ensuring IoT network security. Also, Ismail et al. [15] introduced an ML detection module that utilises the Light Gradient Boosting Machine (LightGBM) algorithm to classify hostile nodes. Based on a comprehensive evaluation of performance metrics, including accuracy, precision, recall, F1-score, processing time, training time, prediction time, computational complexity, and Matthews Correlation Coefficient (MCC), LightGBM has been identified as the optimal choice among the various machine learning algorithms tested on the WSN-DS dataset. These algorithms include Logistic Regression, Complement Naive Bayes, Nearest Centroid, and Stacking.

Several studies investigate intrusion detection in blockchain. These studies highlighted some of the techniques that can be used to effectively detect malicious nodes in blockchain such as PRO-DLBIDCPS which improves feature selection with AHSA and employs ABi-GRNN. DLBC, BEML and LightGBM are other techniques proposed to integrate the ML models to provide better accuracy. While these approaches show promise, further research is needed to address scalability, real-world implementation challenges, and the evolving nature of threats.

## 3 Proposed Multi-Layer Perceptron Model

There is a need for an effective system to identify malicious nodes in blockchain environments. This applied research has developed a novel DL technique to identify and detect malicious nodes in PoS permissionless blockchains. The proposed model detects malicious activities by analysing transaction data, node behaviour, and network traffic using Multi-Layer Perceptron (MLP). It integrates components for data collecting and feature extraction. The model provides a comprehensive solution to improve the security and integrity of PoS blockchain networks by focusing on real-time detection and decision-making, scalability considerations, and security measures, which can meaningfully help protect against malicious actors. The proposed model is shown in Figure1. The proposed model involves the following modules:

- **Data Collection**: This component retrieves data from the blockchain network, including transaction data, block data, network traffic, and metrics related to node behaviour. This data is used as the input for the DL model.
- **Feature Extraction**: The main functionality of this module is to identify and extract important features from the gathered data. Possible features may encompass transaction volume, transaction frequency, block generation time, node reputation scores, and network latency. Feature extraction may also encompass pre-processing procedures like normalisation and scaling.
- **Model Training (MLP)**: MLPs are used to identify rogue nodes in blockchain networks accurately. MLPs are crucial in analysing complex patterns and connections in the data obtained from blockchain transactions, node behaviour, and network interactions. The MLPs are trained using labelled data that represents both

benign and malicious node activities, by utilising multiple variables such as transaction volumes, frequencies, timestamps, and node reputation scores. This MLP-based technique utilises iterative training and validation to identify small irregularities that are suggestive of malicious behaviours, such as double spending or Sybil attacks. This approach allows for pre-emptive actions to reduce risks, ultimately improving the overall security of the blockchain network.

- **Detection Decision**: The detection decision module for blocking malicious nodes and allowing normal nodes is designed to leverage the outputs of the MLP-based detection system. Upon detecting anomalous behaviour indicative of malicious activity, the decision module employs predefined thresholds or rules to determine whether to block or allow the node. Factors such as the severity of detected anomalies, the node's reputation score, and the potential impact on the blockchain network's security are taken into consideration. For normal nodes, the decision module ensures uninterrupted participation in the network by allowing their transactions and interactions to proceed without intervention. This proactive approach enables swift responses to threats while maintaining the integrity and functionality of the blockchain network.

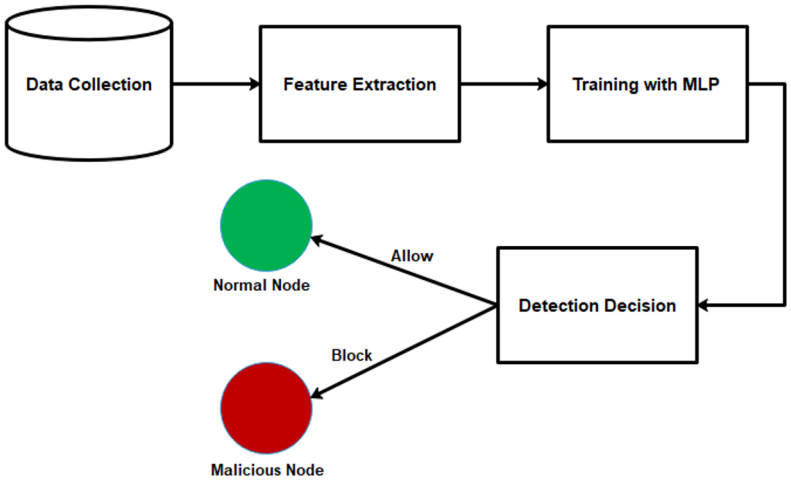


**Fig. 1.** The proposed model using Multi-Layer Perceptron.

## 4 Testing, Validation, Results and Discussions

The implementation of the proposed model consists of two main phases: exploratory data analysis and implementation of the MLP model. This section will highlight the details of these phases as follows.

### 4.1 Exploratory Data Analysis

Exploratory Data Analysis (EDA) in intrusion detection involves examining and analysing data to understand its characteristics, identify patterns, and gain insights into potential security threats or anomalies. The research applied the proof-of-stake blockchain dataset from Kaggle [16]. The dataset's EDA uncovers numerous significant findings. The correlation matrix shown in Figure 2 illustrates the extent of the linear connection between

pairs of numerical characteristics, where higher correlation values indicate greater relationships.

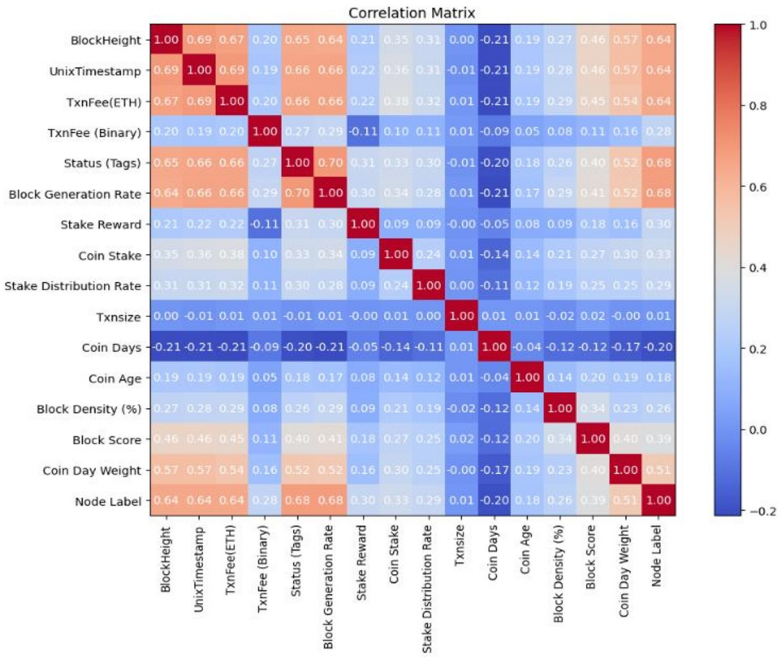


**Fig. 2.** Correlation matrix of numerical features from the dataset.

## 4.2 Implementation of MLP algorithm

MLP is used to detect malicious nodes in the blockchain. This approach involves a sequential model with three dense layers. The initial layer consists of 50 units that employ the rectified linear unit (ReLU) activation function. It receives input shape information obtained from the flattened training data. The second concealed layer is composed of 50 units with the ReLU activation function. The output layer, which consists of 2 units and has a SoftMax activation function, represents the class. The model is created with the Adam optimiser, categorical crossentropy loss function, and accuracy as the evaluation metric. This architecture is specifically built to capture intricate linkages among the input data and precisely categorise defects. The MLP model's output differentiates between malicious nodes and normal/regular nodes inside the blockchain network, offering vital data for network security and integrity. The graphical representation of the MLP model for both training and validation for accuracy and loss values can be seen in Figure 3.

The accuracy and loss values for MLP for both training and validation shown in Figures 3, 4 and 5 demonstrate that how well the model performed during training. This shows that the model achieved an accuracy of about 98% for the training data and about 98% for the validation or testing data. The blue line represents the model training accuracy, whereas the orange line represents the validation test accuracy. The graph also shows a representation of the losses acquired by the model during training and testing. The green line indicates the loss acquired by the model during training, and the orange line indicates the loss acquired by the model during testing. The loss values are acquired at each training step, starting from step 1 to step 8. Loss values mean the losses the model had during training. This shows that the model achieved a loss value of about 0.06% for the training data and the validation or testing data.

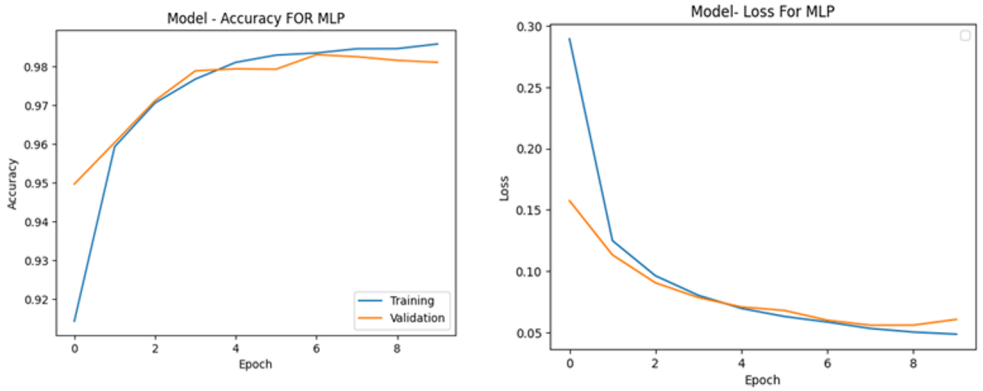


**Fig. 3**. Accuracy and loss values for MLP for both training and validation

Classification_Report For MLP

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Normal_node | 0.99 | 1.00 | 0.99 | 1959 |
| Malicious_Node | 1.00 | 0.99 | 0.99 | 1891 |
| accuracy | | | 0.99 | 3850 |
| macro avg | 0.99 | 0.99 | 0.99 | 3850 |
| weighted avg | 0.99 | 0.99 | 0.99 | 3850 |

**Fig. 4.** Classification values of the MPL model

The classification of the proposed MLP model indicates excellent performance in differentiating between normal and malicious nodes in a network dataset. The model exhibits outstanding accuracy in classification, as evidenced by precision, recall, and F1-score values of about 0.99 for both classes. The model's overall accuracy is given as 0.99, suggesting its high effectiveness in correctly identifying cases from both classes. The macro and weighted average scores provide additional evidence of the model's strength across all classes, as each metric regularly achieves virtually flawless values. These findings indicate that the MLP model demonstrates a high level of dependability in accurately categorising network nodes as either normal or malicious. As a result, it holds significant value as a tool for network security applications.

The experimental results show the performance of the proposed model in differentiating between normal and malicious nodes in a network dataset. The evaluation metrics used, including precision, recall, F1-score, and overall accuracy, indicate the excellent performance of the model. While comparing the proposed model with Sayadi et al. [17], The TPR of the proposed model is 99%, which means it correctly identifies 99% of the malicious nodes in the dataset. In comparison, Sayadi et al. [17] achieved a slightly lower TPR of 98%. This suggests that the proposed system is slightly more capable of accurately identifying malicious nodes. Also, the proposed model's false positive rate (FPR) is significantly lower at 0.00078. A lower FPR indicates that the proposed system can better avoid misclassifying normal nodes as malicious. This is crucial in network security applications, as misclassifying normal nodes as malicious could lead to unnecessary alarms or disruptions.

**Table 1.** Evaluation of the proposed model against models from the literature

| Related Models | True positive Rate | False Positive rate | Accuracy (%) |
|---|---|---|---|
| Sayadi et al. [17] | 99 | 0.78 | 90 |
| Morishima [18] | 98 | 0.05 | 98 |
| **Proposed model** | **99** | **0.00078** | **99** |

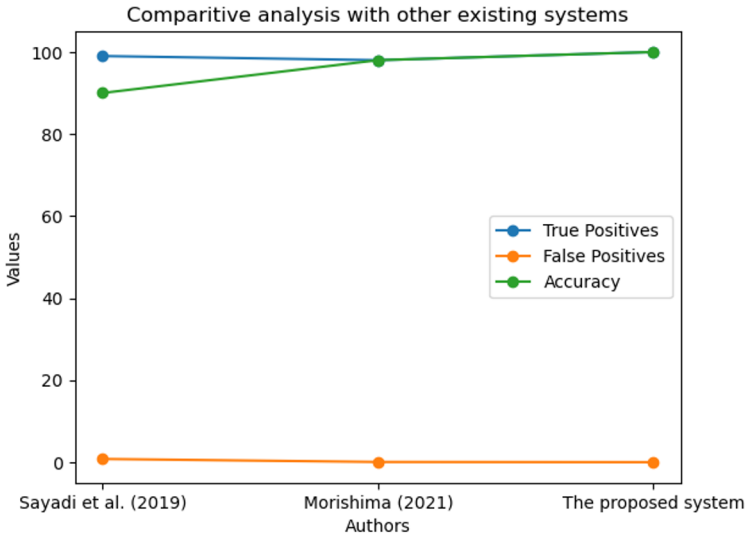


**Fig. 5**. Comparison of the proposed model against models from the literature

In terms of overall accuracy, the proposed model achieves a score of 99%, which is higher than the 90% accuracy reported by Sayadi et al. [17]. This demonstrates that the proposed model performs better in categorising both normal and malicious nodes correctly. Comparing the proposed model to the model presented by Morishima [18], it can be clearly seen that both models have high true positive rates (99% for the proposed model and 98% for Morishima). However, the FPR of the proposed system is significantly lower at 0.00078. This indicates that the proposed model has a better ability to differentiate between normal and malicious nodes without misclassifying normal nodes. The overall accuracy of the proposed model (99%) is also slightly higher than Morishima's [18] model (98%). This suggests that the proposed model outperforms Morishima's model in accurately classifying network nodes.

## 5 Conclusion

This paper proposed and presented a novel approach utilising MLP models for the real-time detection of malicious nodes in PoS blockchain networks. The model showcases exceptional performance, with accuracy scores, precision, recall, and F1-scores consistently near-perfect for both normal and malicious node classifications. By leveraging features extracted from transaction data, node behaviour, and network traffic, the MLP-based model achieves unparalleled accuracy in distinguishing between benign and malicious activities. Comparative analysis with existing works demonstrates the superiority of the proposed model, particularly in its significantly lower false positive rate. These findings emphasise the potential of MLP-based techniques in bolstering the security and integrity of PoS blockchain networks, offering proactive measures against malicious actors while ensuring the trustworthiness of blockchain transactions. For future work of this ongoing research programme, the scalability and applicability of the proposed model in larger and more

complex blockchain ecosystems will be considered, ultimately advancing the resilience of decentralised systems against evolving cybersecurity threats.

## References

1. D. Marbouh, T. Abbasi, F. Maasmi, I. Omar, M. Debe, K. Salahet Arab. J. Sci. Eng, **45** (12), p. 9895-9911(2020)
2. N. T. Anthony, M. Shafik, F. Kurugollu, H. F. Atlam, *Anomaly Detection System for Ethereum Blockchain Using Machine Learning*, 19th International Conference on Manufacturing Research ICMR2022, 5-8 September 2022 Derby, UK, (2022)
3. A. Cetinkaya, H. Ishii, & T. Hayakawa, Entropy, **21** (2), p. 210 (2019)
4. K. Albulayhi, A. Smadi, F. Sheldon, & R. Abercrombie, Sensors, **21** (19), p. 6432 (2021)
5. L. Huang, G. Jia, W. Fang, W. Chen, & W. Zhang, Sensors, **20** (1), p. 221 (2019)
6. B. Kim, K. Kim, B. Shah, F. Chow, & K. Kim, Sensors, **19** (7), p. 1565 (2019)
7. S. Shi, D. He, L. Li, N. Kumar, M. Khan, & K. Choo, Comput. Secur, **97**, p. 101966 (2020)
8. R. Asif, K. Ghanem, & J. Irvine, Sensors, **21** (1), p. 28 (2020)
9. W. She, Q. Liu, T. Zhao, J. Chen, B. Wang, & W. Liu, IEEE Access, **7**, p. 38947-38956 (2019)
10. W. Li, Y. Wang, W. Meng, L. Jin, & C. Su, IEICE Trans. Inf. Syst, **E105.D** (2), p. 272-279 (2022)
11. R. .F Mansour, Scientific Reports, **12**(1), 12937 (2022)
12. S. Hisham, M. Makhtar, A. A. Aziz, Int. J. Adv. Technol. Eng. Explor, **9** (95), 1366 (2022)
13. B. Li, C. Chenli, X. Xu, Y. Shi, T. Jung, *Dlbc: A deep learning-based consensus in blockchains for deep learning services.* arXiv preprint arXiv:1904.07349 (2019)
14. S. Musa Baig, M. U. Javed, A. Almogren, N. Javaid, M. Jamil, Peer-to-Peer Netw. Appl, **16** (6), 2811-2832 (2023)
15. S. Ismail, M. Nouman, D. W. Dawoud, H. Reza, Blockchain Res. Appl, **5** (1)100174 (2023)
16. Kaggle, proof-of-stake blockchain dataset, https://www.kaggle.com/datasets/a9910rut/proofofstake-blockchain-dataset Last accessed [12 May 2024].
17. S. Sayadi, S. Ben Rejeb, Z. Choukair, *Anomaly Detection Model Over Blockchain Electronic Transactions,* 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), 24-28 June 2019, pp. 1-5, (2019)
18. S. Morishim, Comput Electr Eng, **92**, 107087 (2021)